Example P1 analysis and figures

Created by: Michael Van Nuland (Feb. 2017)

This is a step-by-step guide on how to handle the P1 data, conduct the statistical tests, and generate the figures. This guide uses a dataset that is identical to the one on Blackboard in every way - except for the actual data. That mean you can run this exact code with the real dataset, just make sure in STEP 2 that you load in the correct file (i.e., NOT the example file). Keep in mind your specific results and figures will look different than the output here. If you get stuck, have a look at the more detailed R tutorial posted on BB.

STEP 1. Download the "P1_data_example.csv" file from blackboard and **save it to your desktop**. (Do the same for the real file once I have all your data to upload it).

STEP 2. Open RStudio and set the working directory to your desktop. Do this by going to Sessions (top menu bar) -> Set working directory -> Choose directory -> Desktop

STEP 3. Read in the data file and assign it to a "P1_data" object. Because you've set your working directory to your desktop, and that's where you downloaded the data to, there should be no issues. If you get an error, go back and make sure you did STEP 1 and STEP 2 succesfully.

```
P1_data <- read.csv(file="P1_data_EXAMPLE.csv")    # Again, the real data will be "P1_data.csv"
```

STEP 4. Take a look at the first lines of the data to get oriented. Notice how the species are grouped whether they are adapted to low or high nitrogen. We then have the nitrogen treatments, water treatments, initial data (t0), final data (t26), and biomass data.

```
head(P1_data)
```

```
##     Group    Euc_species N_adapted N_treatment H2O_treatment Height_t0
## 1 Group 1 E. brookeriana  High_spp      High_N           Wet        22
## 2 Group 1 E. brookeriana  High_spp      High_N           Dry        21
## 3 Group 2  E. pauciflora  High_spp      High_N           Wet        16
## 4 Group 2  E. pauciflora  High_spp      High_N           Dry        18
## 5 Group 6 E. brookeriana  High_spp       Low_N           Wet        20
## 6 Group 6 E. brookeriana  High_spp       Low_N           Dry        22
##   LeafNumber_t0 Height_t26 LeafNumber_t26 Aboveground_biomass
## 1             4         38              6                70.0
## 2             4         34              6                66.3
## 3             3         22              7                61.7
## 4             3         20              7                61.5
## 5             4         34              8                79.9
## 6             4         36              8                78.5
##   Belowground_biomass
## 1                28.3
## 2                27.6
## 3                24.0
## 4                22.8
## 5                39.1
## 6                36.1
```

STEP 5. Calculate relative growth rate (RGR) for height, root-to-shoot ratio, and add both to the P1_data object. NOTE: you can do the same for the leaf number data, just make a different object than "RGR" (maybe something like "Leaf_RGR")

```
RGR <- ((log(P1_data$Height_t26) - log(P1_data$Height_t0))/(26 - 0))

Root_shoot <- (P1_data$Aboveground_biomass/P1_data$Belowground_biomass)
```

```
P1_data <- cbind(P1_data, RGR, Root_shoot)

head(P1_data)  # Double check that there is a new RGR column at the end of the file
```

```
##     Group    Euc_species N_adapted N_treatment H2O_treatment Height_t0
## 1 Group 1 E. brookeriana  High_spp      High_N           Wet        22
## 2 Group 1 E. brookeriana  High_spp      High_N           Dry        21
## 3 Group 2  E. pauciflora  High_spp      High_N           Wet        16
## 4 Group 2  E. pauciflora  High_spp      High_N           Dry        18
## 5 Group 6 E. brookeriana  High_spp       Low_N           Wet        20
## 6 Group 6 E. brookeriana  High_spp       Low_N           Dry        22
##   LeafNumber_t0 Height_t26 LeafNumber_t26 Aboveground_biomass
## 1             4         38              6                70.0
## 2             4         34              6                66.3
## 3             3         22              7                61.7
## 4             3         20              7                61.5
## 5             4         34              8                79.9
## 6             4         36              8                78.5
##   Belowground_biomass         RGR Root_shoot
## 1                28.3 0.021020912   2.473498
## 2                27.6 0.018532234   2.402174
## 3                24.0 0.012248220   2.570833
## 4                22.8 0.004052328   2.697368
## 5                39.1 0.020408779   2.043478
## 6                36.1 0.018941403   2.174515
```

STEP 6. Hypothesis 1 relates to how the different N adapted groups respond to nitrogen treatments (*the hypothesis you write for your papers must be more specific than this*). We'll test this with a two-way ANOVA (consult the other tutorial for more details). NOTE: the following example uses Aboveground biomass, but you should explore patterns in all the relavent traits (RGR, Aboveground and Belowground biomass, etc).

```
Hyp1 <- lm(Aboveground_biomass ~ N_adapted*N_treatment, data = P1_data)

anova(Hyp1)
```

```
## Analysis of Variance Table
##
## Response: Aboveground_biomass
##                       Df  Sum Sq Mean Sq F value    Pr(>F)
## N_adapted              1 1620.00 1620.00 44.3440 5.490e-06 ***
## N_treatment            1 1326.67 1326.67 36.3148 1.763e-05 ***
## N_adapted:N_treatment  1  118.41  118.41  3.2411   0.09069 .
## Residuals             16  584.52   36.53
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

To see which groups and treatments are different from one another, run a Tukey's post-hoc test. Comparisons that are different (left-most column) will have significant p-vales (right-most column).

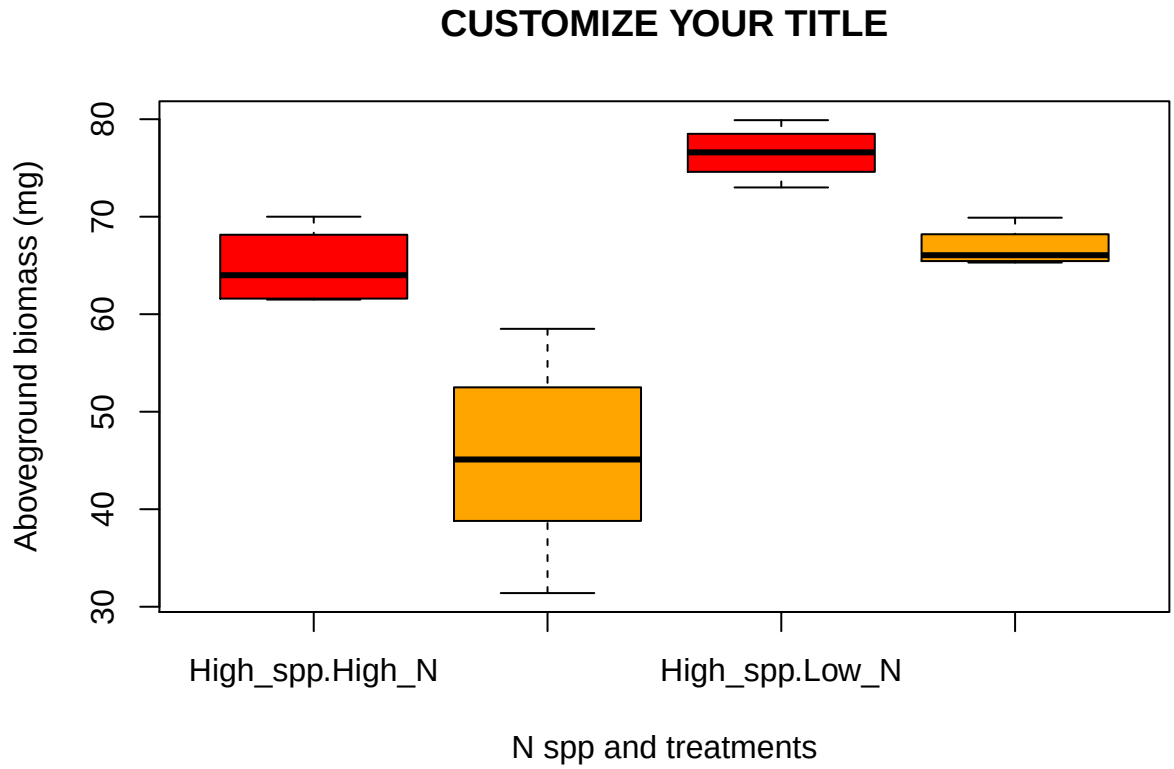
```
TukeyHSD(aov(Aboveground_biomass ~ N_adapted*N_treatment, data = P1_data))
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Aboveground_biomass ~ N_adapted * N_treatment, data = P1_data)
##
```

```
## $N_adapted
##                   diff       lwr       upr   p adj
## Low_spp-High_spp  -18 -23.73023 -12.26977 5.5e-06
##
## $N_treatment
##               diff      lwr      upr    p adj
## Low_N-High_N 15.96 10.22977 21.69023 2.22e-05
##
## $`N_adapted:N_treatment`
##                                         diff         lwr       upr     p adj
## Low_spp:High_N-High_spp:High_N -19.641667 -30.8040148 -8.479319 0.0006369
## High_spp:Low_N-High_spp:High_N  11.658333   0.4959852 22.820681 0.0391119
## Low_spp:Low_N-High_spp:High_N    1.950000 -10.2777397 14.177740 0.9674275
## High_spp:Low_N-Low_spp:High_N   31.300000  21.3160923 41.283908 0.0000007
## Low_spp:Low_N-Low_spp:High_N    21.591667  10.4293186 32.754015 0.0002399
## Low_spp:Low_N-High_spp:Low_N    -9.708333 -20.8706814  1.454015 0.1001313
```

Finally, make a boxplot figure to show the results of Hypothesis 1. Feel free to customize the graph with different colors, etc to make any patterns more clear. To save the file, go to the bottom right where the graph appears and export to a PDF or jpeg. Be sure to make the graph wider/larger to show all the treatments on the x-axis.

```
boxplot(Aboveground_biomass ~ N_adapted*N_treatment, data = P1_data,
      col=(c("red","orange")),   # Red = high N species, Orange = low N species
      main="CUSTOMIZE YOUR TITLE", xlab="N spp and treatments", ylab="Aboveground biomass (mg)")
```



STEP 7. Hypothesis 2 relates to how the nitrogen and water treatments interact to influence plant responses (*the hypothesis you write for your papers must be more specific than this*). We're interested in also seeing if being adapted to low versus high N affects how the treatments interact. First we'll separate the data into the high and low N groups.

```
HighSpp_P1_data <- P1_data[1:10, 1:13]  # Take the first 10 rows that are only the High N adapted speci
LowSpp_P1_data <- P1_data[11:20, 1:13]  # Ditto for the last 10 rows and Low N adapted species
```

Now, we can test how nitrogen and water treatments interact for the two N adapted species groups. Same drill as before: two-way ANOVA, Tukey's post-hoc, and boxplot figure.

```
Hyp2.1 <- lm(Aboveground_biomass ~ N_treatment*H2O_treatment, data = HighSpp_P1_data)

anova(Hyp2.1)
```

```
## Analysis of Variance Table
##
## Response: Aboveground_biomass
##                            Df Sum Sq Mean Sq F value   Pr(>F)
## N_treatment                 1 326.20  326.20 25.9880 0.002226 **
## H2O_treatment               1   8.28    8.28  0.6597 0.447672
## N_treatment:H2O_treatment   1   0.03    0.03  0.0022 0.963755
## Residuals                   6  75.31   12.55
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

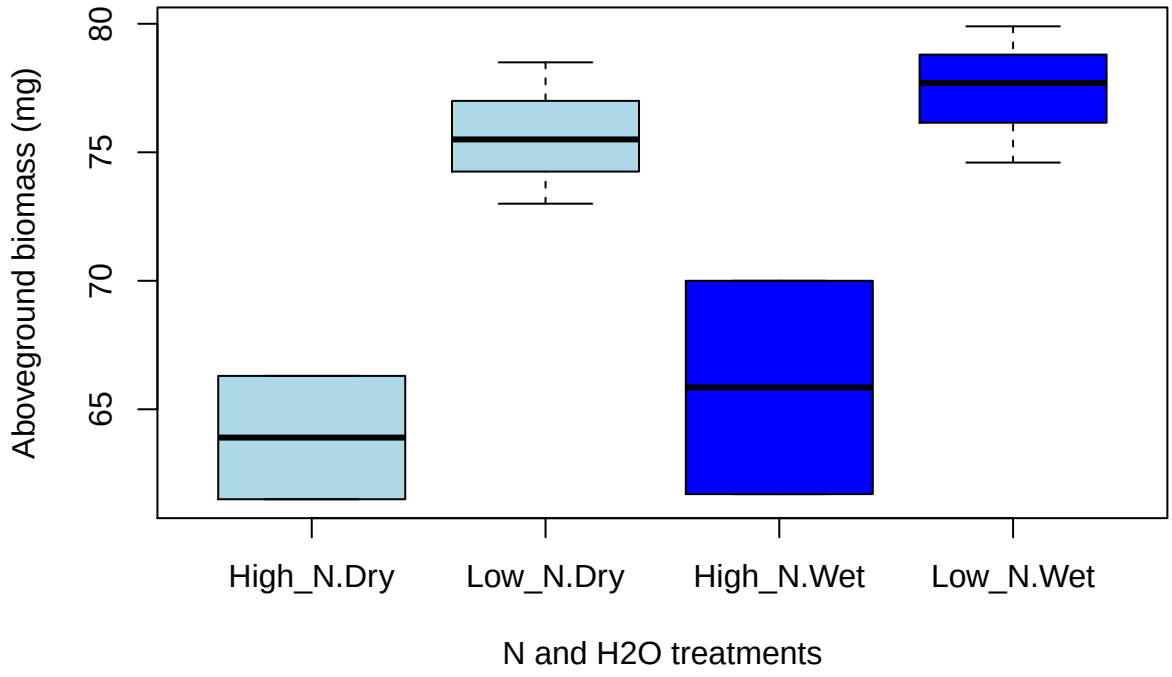
Bring on that Tukey's

```
TukeyHSD(aov(Aboveground_biomass ~ N_treatment*H2O_treatment, data = HighSpp_P1_data))
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##
## Fit: aov(formula = Aboveground_biomass ~ N_treatment * H2O_treatment, data = HighSpp_P1_data)
##
## $N_treatment
##                  diff      lwr      upr     p adj
## Low_N-High_N 11.65833 6.062456 17.25421 0.0022261
##
## $H2O_treatment
##         diff       lwr      upr     p adj
## Wet-Dry 1.82 -3.662818 7.302818 0.4476715
##
## $`N_treatment:H2O_treatment`
##                             diff         lwr       upr     p adj
## Low_N:Dry-High_N:Dry  11.766667   0.5708511 22.962482 0.0408556
## High_N:Wet-High_N:Dry  1.950000 -10.3144015 14.214402 0.9430917
## Low_N:Wet-High_N:Dry  13.500000   2.3041844 24.695816 0.0226203
## High_N:Wet-Low_N:Dry  -9.816667 -21.0124823  1.379149 0.0824904
## Low_N:Wet-Low_N:Dry    1.733333  -8.2805086 11.747175 0.9286671
## Low_N:Wet-High_N:Wet  11.550000   0.3541844 22.745816 0.0440931
```

And finally, a boxplot with different colors to visualize the water treatments.

```
boxplot(Aboveground_biomass ~ N_treatment*H2O_treatment, data = HighSpp_P1_data,
      col=(c("lightblue","lightblue", "blue", "blue")),   # lightblue = dry, blue = wet
      main="CUSTOMIZE YOUR TITLE FOR HIGH N SPP RESPONSE", xlab="N and H2O treatments", ylab="Abovegrou
```



You should do the exact same process for Low N adapted species by changing which data you use (should be the LowSpp_P1_data).

STEP 8. Add your specific hypotheses, statistics (F values and p-values), and figures to your papers. Remember, you should explore patterns in all the different traits, but that doesnt mean you should include all the stats or every possible figure. It's best to examine all the results and then "find the story" in the data - what are the clearest patterns, what things are surprising, what things are novel, what things are important in the broader context of nitrogen pollution, etc. Although each group has the same dataset to analyze to test the same two hypotheses, we expect that your papers will differ in the specific traits and results you report.

In your results, be precise and mention the effect size that different treatments had.

Here's a **bad example**: "High nitrogen spp. responded differently to fertilizer treatments than low nitrogen spp."

Here's a **good example**: "High nitrogen spp. grew 35% taller than low nitrogen spp. in the high fertilizer treatment."